AMENDMENTS TO THE SPECIFICATION

At page 4, lines 9-17

B1

In developing the systems and methods described herein, it was postulated that methods which attempt to describe the entirety of the chemical space of interest using only a few hundred molecules simply do not sample the smaller area of chemical space related to the particular property of interest well enough to be able to make good predictions. Methods such as the hashkey and nearest neighbor approaches exemplify such approaches. In developing a model for protein binding prediction which does not suffer from this drawback, a model development method having wide applicability to activity prediction has been created. Applying the method to the protein binding problem has further resulted in an improved protein binding prediction model.